Amended Claims

1. (Original) A terminal for receiving and processing a multimedia data bitstream, comprising:

a terminal manager;

a composition engine;

a plurality of content decoders; and

a presentation engine; wherein:

said content decoders recover and decode multimedia objects from respective elementary streams of the bitstream;

said multimedia objects comprising at least one of video objects and audio objects for presentation in a multimedia scene;

said composition engine recovers scene description information from the bitstream that defines specific ones of the recovered multimedia objects that are to be provided in the multimedia scene, and characteristics of the recovered multimedia objects in the multimedia scene;

said terminal manager recovers object descriptor information from the bitstream that associates said recovered multimedia objects with respective ones of said elementary streams, and provides the recovered object descriptor information to said composition engine;

said composition engine is responsive to said recovered object descriptor information provided thereto and said recovered scene description information for creating a list of said specific ones of the recovered multimedia objects that are to be displayed in said multimedia scene; and

said presentation engine obtains said list from said composition engine, and, in response thereto, retrieves the corresponding decoded multimedia objects from said content decoders to provide data corresponding to the multimedia scene to an output device.

2. (Original) The terminal of claim 1, wherein:
said composition engine and said presentation engine have separate control threads.

3. (Original) The terminal of claim 2, wherein:
said separate control threads allow the presentation engine to begin retrieving the corresponding decoded multimedia objects while the composition engine recovers additional scene description information from the bitstream and/or processes additional object descriptor information provided thereto.

4. (Original) The terminal of claim 1, wherein:
said content decoders, presentation engine and composition engine have separate control threads.

5. (Original) The terminal of claim 1, wherein:
said characteristics of the recovered multimedia objects in the multimedia scene include positions of said specific ones of the recovered multimedia objects in said multimedia scene.

6. (Original) The terminal of claim 1, wherein:
said recovered scene description information is provided according to a Binary Format for Scenes (BIFS) language

7. (Original) The terminal of claim 1, wherein:
said multimedia data bitstream is provided according to an MPEG-4 standard.

8. (Original) The terminal of claim 1, wherein:
said composition engine maintains scene graph information of a composition of said multimedia scene in response to said recovered object descriptor information provided thereto and said recovered scene description information for use in creating said

list.

9. (Original) The terminal of claim 8, wherein:

said composition engine updates the scene graph information, and said list, as required, for successive multimedia scenes in response to subsequent recovered scene description information from the bitstream.

10. (Original) The terminal of claim 8, wherein:

said terminal manager is responsive to user input events at a user interface for providing corresponding data to said composition engine for modifying said scene graph, and said list, as required.

11. (Original) The terminal of claim 1, wherein:

said composition engine provides said list to said presentation engine according to a specified presentation rate.

12. (Original) The terminal of claim 1, wherein said multimedia objects comprise video and audio objects for presentation in the multimedia scene, further comprising:

video and audio buffers for buffering the video and audio objects, respectively, prior to presentation;

wherein said presentation engine reads objects from said list and provides them to the appropriate one of said video and audio buffers.

13. (Original) A terminal for receiving and processing a multimedia data bitstream, comprising:

decoding means for recovering and decoding multimedia objects from respective elementary streams of the bitstream;

said multimedia objects comprising at least one of video objects and audio objects for presentation in a multimedia scene;

composing means for recovering scene description information from the bitstream that defines specific ones of the recovered multimedia objects that are to be provided in the multimedia scene, and characteristics of the recovered multimedia objects in the multimedia scene;

managing means for recovering object descriptor information from the bitstream that associates said recovered multimedia objects with respective ones of said elementary streams, and providing the recovered object descriptor information to said composing means;

said composing means being responsive to said recovered object descriptor information provided thereto and said recovered scene description information for creating a list of said specific ones of the recovered multimedia objects that are to be displayed in said multimedia scene; and

presenting means for obtaining said list from said composing means, and, in response thereto, retrieving the corresponding decoded multimedia objects from said decoding means to provide data corresponding to the multimedia scene to an output device.

14. (Original) A method for receiving and processing a multimedia data bitstream at a terminal, comprising the steps of:

recovering and decoding multimedia objects from respective elementary streams of the bitstream at respective content decoders;

said multimedia objects comprising at least one of video and audio objects for presentation in a multimedia scene;

recovering scene description information from the bitstream that defines specific ones of the recovered multimedia objects that are to be provided in the multimedia scene, and characteristics of the recovered multimedia objects in the multimedia scene;

recovering object descriptor information from the bitstream

that associates said recovered multimedia objects with respective ones of said elementary streams;

creating a list of said specific ones of the recovered multimedia objects that are to be displayed in said multimedia scene in response to said recovered object descriptor information and said recovered scene description information; and

retrieving the corresponding decoded multimedia objects in response to the list to provide data corresponding to the multimedia scene to an output device.

15. (Original) The method of claim 14, wherein:

said recovering steps are performed using control threads that are separate from said retrieving step.

16. (Original) The method claim 15, wherein:

said separate control threads allow the retrieving of the decoded multimedia objects to begin while the recovering of additional scene description information and/or the recovering of additional object descriptor information occurs.

17. (Original) The method of claim 14, wherein:

said creating step is performed using a control thread that is separate from said retrieving step.

18. (Original) The method of claim 14, wherein:

said recovering steps and said creating step are performed using control threads that are separate from said retrieving step.

19. (New) The terminal of claim 1, further comprising:

an adaptation layer for receiving the bitstream;

wherein:

the content decoders receive the respective elementary

streams from the adaptation layer;

the composition engine receives said scene description information from the adaptation layer; and

the terminal manager receives said object descriptor information from the adaptation layer.

20. (New) The terminal of claim 13, further comprising:

an adaptation layer for receiving the bitstream;

wherein:

the decoding means receive the respective elementary streams from the adaptation layer;

the composing means receives said scene description information from the adaptation layer; and

the managing means receives said object descriptor information from the adaptation layer.

21. (New) The method of claim 14, wherein:

an adaptation layer for receiving the bitstream is provided;

the content decoders receive the respective elementary streams from the adaptation layer;

said scene description information is recovered from the adaptation layer; and

said object descriptor information is recovered from the adaptation layer.